RCA PF030181AD

CITED BY APPLICANT

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 May 2001 (17.05.2001)

PCT

(10) International Publication Number
WO 01/35271 A2

(51) International Patent Classification[7]: G06F 17/30

(21) International Application Number: PCT/US00/29819

(22) International Filing Date: 30 October 2000 (30.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/162,465 29 October 1999 (29.10.1999) US

(71) Applicant: GIGABEAT, INC. [US/US]; 345 California Avenue, Suite 3, Palo Alto, CA 94306 (US).

(72) Inventor: JANNINK, Jan, F.; 2041 Sharon Road, Menlo Park, CA 94025 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *Without international search report and to be republished upon receipt of that report.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: SYSTEMS AND METHODS FOR VISUALIZATION OF DATA SETS CONTAINING INTERRELATED OBJECTS

(57) Abstract: A visualization of a set of related data items is accomplished by ranking a plurality of secondary data items with respect to their relationship to a principal data item. The set of related data items may be stored in any form, including, for example, as a relational database, an object database, or an XML document. The related items display names can be formatted and ranked according to the properties of the item, so as to avoid duplication and information overload for the user. The ranked data can be displayed along curved segments in a graphical chart, so as to allow convenient visualization of the data and the data's relationship to the principal item that may be prominently displayed at the center of the chart. The visualization of data sets containing a large number of items from data sets is enabled, while respecting the size constraints of the display and maintaining an optimal level of user comprehension. The visualization has the property that subsets of items in a data set relate to each other, and the relationships between items have a value associated with each other. Local rankings of the relationships between terms are established, by ranking items that relate to another. Next, a visualization is generated in the form of a graphical chart, by presenting results separately for each item in a predetermined data set and adjusting the presentation to avoid information overlap and overload. Separate charts for each item of the data set can be presented. Further, closeness is expressed by proximity of items along a curved segment, which are connected or which emanate from a selected position. To enhance the clarity of the visualization to avoid information overlap and overload, the items related to a particular item are grouped by strength of affinity. Each related item is individually spaced on the affinity chart, with each item being placed in a non-overlapping position. Items with large numbers of related items are presented with multiple affinity charts. In the case of multiple affinity charts, a first affinity chart visualizes a set of most strongly related items. Next or subsequent related affinity charts visualize less strongly related items. Color and shading gradations and curve thickness gradations are selectively employed to emphasize the curve's role in conveying affinity strength. Items are placed so they do not overlap or crowd each other. Arbitrarily large data sets are visualized using low and localized computational resources.

WO 01/35271 A2

# SYSTEMS AND METHODS FOR VISUALIZATION OF DATA SETS CONTAINING INTERRELATED OBJECTS

## *RELATED APPLICATIONS*

The present application claims priority to Provisional Application No. 60/162,465 filed October 29, 1999. The present application also relates to co-pending U.S. Application Serial Number 09/574,108, filed May 19, 2000. Each of the above referenced applications is incorporated herein by reference in its entirety.

## *BACKGROUND*

### Field of the Invention

The present invention relates to visualization systems and methods and more particularly to systems and methods for enabling visualization of data sets containing large numbers of objects.

### Description of the Related Art

U.S. Patent No. 5,966,139, granted to Anupam et al. and assigned to Lucent Technologies Inc., relates to scalable data segmentation and visualization. In particular, the '139 patent describes visualization of relatively large amounts of data in a limited display space, including segmentation of data, mapping of segments to a node within a display space based upon a specified nodal layout, texture mapping each graphical display to its node, and displaying the data at each node. The visualization scheme enables a user to map a relation to a specification of an n-dimensional display, by designating how attributes are to be used.

However, the visualization system of the '139 patent fails to represent data sets having two properties, namely that subsets of items in a data set relate to each other, and the relationships between items have an affinity value associated with each other. Furthermore, a system according to the '139 patent does not simplify the display of data to the user by presenting multiple screens of only the amount of data that can be comfortably fit on a screen.

## *SUMMARY OF THE INVENTION*

The present invention discloses methods and systems for the visualization of data sets containing a large number of interrelated objects (or items) from data sets having two properties, namely that subsets of items in a data set relate to each other, and the relationships between items have a value associated with each other. In particular, according to an embodiment of the present invention, a compact, easily understood, and easily navigable visual representation of objects in a data set can be achieved.

As a preparatory step, local rankings of the relationships between items are established, by ranking the items i that relate to each item j, and ranking all items k to which item j relates, thereby ranking the affinity to each item j of item sets i and k. Next according to an embodiment of the present invention, a visualization can be generated, by presenting results separately for each item in a predetermined data set and adjusting the presentation to avoid information overlap and overload. Separate representation of each item of the data set can be accomplished by generating an affinity chart for each item j in the data set, to display items closely related to selected item j, with item j being placed prominently in the affinity chart, and placing items which are more strongly related to item j closer to item j. Further, closeness is expressed along curves or shaped segments which may be completely or partially straight, which are connected or which are emanating from item j's position.

According to one embodiment of the present invention, continuous curves including but not limited to spiral segments, are employed to connect items relating to item j at different intensity levels. To adjust the visualization to avoid information overlap and overload, the items related to a particular item j are grouped by strength of affinity. Each related item is individually spaced on the affinity chart, with each item being placed in a non-overlapping position. Items with large numbers of related items are presented with multiple affinity charts. In the case of multiple affinity charts, a first affinity chart visualizes a set of most strongly related items. Next or subsequent related affinity charts visualize less strongly related items. According to an embodiment, curves can be used to represent the relationship of items related to a particular item positioned at a starting point for the curve. Distance along the affinity curve represents

the strength of the affinity to the item at the starting point of the curve. Color and shading gradations and curve thickness gradations are selectively employed to emphasize the curve's role in conveying affinity strength. Items are placed so they do not overlap or crowd each other. Arbitrarily large data sets are visualized using low and localized computational resources.

## *BRIEF DESCRIPTION OF THE DRAWINGS*

Fig. 1 is an affinity chart according to an embodiment of the present invention.

Fig. 2 is an affinity chart of other related items for the affinity chart of Fig. 1.

Fig. 3 is a flowchart illustrating one method of visualizing large interrelated data sets.

Fig. 4 is a data flow diagram showing the flow of data through a system for visualizing large interrelated data sets.

Fig. 5 depicts a system for visualizing large interrelated data sets.

Fig. 6 illustrates a type of database structure that can be used as a source for the rankings used by a system visualizing large interrelated data sets.

Fig. 7 is another affinity chart, according to the present invention.

Fig. 8 is another affinity chart of other related items for the affinity chart of Fig. 7.

## *DETAILED DESCRIPTION*

Visualization of a data set refers to the use of various techniques to convey the overall structure of information by visual means. In particular, visual cues can be used to represent relations between objects. Visual cues can include, for example, using segments of a curve to represent the affinity (or strength) of the relationship between objects, and using gradations of the width and color of the curves to represent the intensity of the affinity relationships between the objects.

Referring now to Fig. 1, there is shown a diagram of most closely related items according to the present invention. In particular, Fig. 1 shows an affinity chart 129 including first and second affinity curves 130 and 140 including a principal item 131 and first and second pluralities of related items 132 and 142. Each related item 132,

142 includes a navigational link 133, 143 respectively and a search link 134, 144 respectively. Adjacent to the principal item 131 and at one end of a selected strings of related items 132 and 142 are respective first sequence element 135 and 145, and adjacent to a last item of the selected strings of related items 132, 142 is a second sequence element 136, 146 which provides a link to a supplementary affinity chart for more remotely relevant strings of related items 132, 142.

In an embodiment, the affinity chart 129 may consist of a single list of textual or graphical items and associated links. The principal item 131, related items 132 and 142, first sequence elements 135, 145, and second sequence elements 136, 146 may all appear as items of the list. In this list, each related item may appear associated with a navigational link 133, 143, and a search link 134, 144. Such an affinity chart may be required when the display in which it is presented only accommodates lists.

Referring now to Fig. 2, there is shown a diagram of other lesser related items to principal item 131 than those shown in Fig. 1, according to the present invention. In particular, Fig. 2 shows an affinity chart 249 representing those items that would have been reached as a result of selecting first sequence element 135 in Fig. 1. Affinity chart 249 in Fig. 2 includes first and second affinity curves 250 and 260 including a principal item 131 and first and second pluralities of related items 252 and 262 that are lesser related to principal item 131 than related items 132 and 142 shown in Fig. 1. Each related item 252, 262 includes a navigational link 253, 263 respectively and a search link 254, 264 respectively. Adjacent to the principal item 131 and at one end of a selected strings of related items 252 and 262 are respective first sequence elements 255 and 265, which provide a link to the affinity chart of more strongly related items, and adjacent to a last item of the selected strings of related items 252, 262 is a second sequence element 256, 266 which provides a link to a supplementary affinity chart for more remotely relevant strings of related items 252, 262. By selecting the first sequence element 255 or 265, the user can navigate to reach the affinity curves 240 or 250 to view a string of relevant items as shown in Fig. 1.

Fig. 3 is a flow chart depicting a method for visualization of large interrelated data sets, according to the present invention. According to this method, the visualization of data sets containing a large number of items from data sets having two

properties, namely that subsets of items in a data set relate to each other, and the relationships between items have a value associated with each other, is enabled.

In an embodiment of the invention, a method of visualization of large interrelated data sets can include an information structuring phase 305, a chart layout phase 310, and an information linking phase 315. In information structuring phase 305, the relationships between objects in a data set and the intensity of those relationships can be computed. As a preparatory step in the information structuring phase, an item from the data set can be selected in step 320. Local rankings of the relationships between items can be established in step 324, by ranking for each selected item j the items i that relate to that item j, and then ranking all items k to which item j relates, thereby ranking the affinity for each item j to item sets i and k. As a result, the rankings of the related items are relative (or local) to item j, but are not a universal measurement of the importance of the item. To determine how an item relates to another item, the strength of the relationship between the items can be computed using any combination of objective or subjective criteria, or a combination of both. A value can be associated with each criterion, and a plurality of these values can be reduced to a single value (i.e. an affinity value), by, for example, adding them all together and then normalizing the value.

Referring back to Fig. 1, principal node 131 ("The Beatles") would have item set i and item set k related to it. Spiral curve 130 depicts item set i (i.e. those items that are related to principal node 131), while spiral curve 140 depicts item set k (i.e. those items to which principal node 131 relates). To determine both item set i and item set k in Fig. 1, both objective and subjective criteria related to The Beatles could be used. Objective criteria used to determine the relationships between various musical bands could include, for example, the era in which the band played (e.g. 1960s), and the genre of the music (e.g. rock, British Invasion, pop). Subjective criteria could include, for example, how well liked the band is based on feedback from users, and how often two bands appear together in radio station play lists.

In an embodiment, a single affinity value can be determined from the various criteria to represent the strength of the relationship between objects, and the related items to the selected item can be ranked in step 324 by the affinity values associated

with each related item. Once ranked, those related items can be clustered in step 328. Clustering refers to the process of appropriately grouping the ranked objects. For example, an arbitrary number (e.g. ten) of the most closely related items can be chosen as a cluster. Once clustered, the number of affinity charts can be computed in step 332. Thus, using the above example, if twenty-eight related items exist, step 332 can result in a computation of three affinity charts needing to be generated, two of which would have ten items and the third having eight items. The first ten would be the most closely related to the principal item, the next ten would be the next most closely related, and the last eight would be the next most closely related.

Next in an embodiment of the invention, chart layout phase 310 can cause the organization and placement of the relevant subsets of the objects in charts or graphs that can be used to display the relationships between the objects. The complete set of charts constitutes a virtual map of the data set.

Chart layout phase 310 can begin by positioning the selected item as a principal node in a chart in step 338. This can include the placement of a hyperlink for the selected item in a specified position of the chart. Next a visualization can be generated, by presenting results separately for each item in a predetermined data set and adjusting the presentation to avoid information overlap and overload. In an embodiment, separate presentation for each item of the data set can be accomplished by generating an affinity curve in step 342 for each item j in the data set, to display items closely related to selected item j, with item j being placed prominently in the affinity chart, and placing items which are more strongly related to j closer to j. Further, closeness can be expressed along curves or shaped segments which may be completely or partially straight, and which may be connected or may emanate from j's position.

According to an embodiment of the present invention, continuous curves including but not limited to spiral segments, can be employed to connect items relating to j at different intensity levels. In step 346, a related item can be selected and in step 350, the size required for that related item can be determined. In step 354, the related item can be individually spaced on the affinity chart by its rank, with each item being placed in a non-overlapping position by allowing sufficient vertical and horizontal displacement in step 358. A determination can be made in step 360 of whether any

more related items need to be placed on the affinity curve. If so, control returns to step 346 where the next item is selected. If there are no more related items to process, the color and size gradients of the curve can be adjusted in step 362 to emphasize the affinity between the items.

A selected item with a large number of related items can be presented with multiple affinity charts. In the case of multiple affinity charts, a first affinity chart can be used to visualize a set of most strongly related items. Next or subsequent related affinity charts can be used to visualize less strongly related items.

According to an embodiment, arbitrarily large data sets can be visualized using low and local computational resources. During information linking phase 315, the sequence of affinity charts for a selected item can be hyperlinked in step 366. Each related item can then be linked to its own chart in step 370. Once the selected item and its associated affinity curves from the data set have been hyperlinked, navigation by the user can occur by the user clicking to connect to a selected related affinity chart. Further, as a result of information linking phase 315, each item may have separate features that can be activated by the user clicking on those features.

Upon completion of information linking phase 315, a determination can be made in step 381 of whether more charts need to be generated for the selected item. If so, control can pass to step 338 where the selected item can be placed in a new chart. If no further charts are needed for selected items, a determination can be made in step 386 of whether more items in the data set need charts generated for them. If so, control can pass to step 320 where a determination is made of a new selected item and the process repeats.

Fig. 4 is a data flow diagram showing the flow of data through a system for visualizing interrelated data sets, according to an embodiment of the present invention. In the system shown in Fig. 4, a user 405 of a computer 412 containing a client browser 410 (as is well known and understood in the art) can send a request 415 via a distributed computer network (such as the Internet) to a web server 420. Web server 420 can process the request and ultimately provide the interface for user 405, including the ability to navigate visualized data sets.

Upon receiving the request, web server 420 can process the request and send the request on to chart server 425. A determination can then be made in step 430 of whether the particular chart corresponding to the request had previously been cached. This could have occurred if a prior request had been made for the same data. If so, the necessary information can be provided from chart cache 435. By way of example only, the information could consist of charts and hyperlinked data 440.

If the particular chart corresponding to the request had not previously been cached, the request can be passed on to affinity server 445. Affinity server 445 can have available to it object affinity tables 450 that contain information regarding the items in the data set, including, for example, the fields detailed in the affinity tables shown in Fig. 6. After receiving a request, affinity server 445 can access object affinity tables 450 for the information needed to structure the chart corresponding to the request, according to information structuring phase 305 described with respect to Fig. 3.

After being retrieved from chart cache 435 or generated by affinity server 445, the charts and hyperlinked data 440 can be passed to chart server 425 where the visualized data set can be graphically laid out and linked together. Upon completion, chart server 425 can pass the graphically laid out and linked visualized data set to web server 420, which, in turn, can pass it on to client browser 410 being used by user 405. As a result, the visualized data set can be displayed to user 405.

Fig. 5 depicts an embodiment of a system according to the present invention for visualizing large interrelated data sets. As shown in Fig. 5, web server 505 can receive image request 507 for a particular artist. Upon determining that image request 507 is to generate a visualization of data from a data set, that request can be passed to chart server 510. Chart server 510 can then determine whether the necessary information to fulfill the request (i.e. the affinity chart) has already been calculated and cached. If so, image map 509 can be retrieved by chart server 510 from cache 515 and returned to web server 505 for display to the user. The retrieval of the bit map can then cause bitmap request 511, which can then be followed by the retrieval of the associated bitmaps from cache 515 and the display of bitmap 513 to the user.

If the information needed to fulfill the request has not been cached, chart server 510 can fulfill the request by retrieving information about the requested item from database 550. In particular, database 550 can contain data objects 555 that contain a distinct affinity value (A) between the requested item (O1) and each related item (O2). In response to a query by cache server 510, one or more data objects 555 for the requested item can be retrieved from database 550 and utilized by affinity server 520 to generate the necessary information (including, for example, image maps and bitmaps) for the selected item. Affinity server 520 can utilize an affinity ranking means 540 to rank the related items by their affinity values, and an image layout program 545 and image layout interpreter 535 to assemble the images. Affinity server 520 can also utilize image map generator 525 and bitmap renderer 530 to generate the actual image map and bitmaps. Once generated, the image map and bitmaps can be placed in cache 515 and transferred to web server 505. Web sever 505 can then present to the user the visualization of the interrelated data associated with the requested item.

Fig. 6 illustrates a type of database structure that can be utilized in the process of visualizing large interrelated data sets, according to an embodiment of the present invention. Within a particular data set, each item can be identified within a table 605 in a database by an object identifier (OID) 610 and an object name 615 (that may also contain object properties). The OID associated with each item is a primary key (PK) 620. Two OIDs from table 605 serve as primary key 620 for the affinity strength table 625. These two OIDs are labeled as foreign keys (e.g. FK1 and FK2 in field 640) of the second affinity strength table 625 within the database. Table 625 can be accessed by fixing the OID values to obtain the affinity value 635 between two items identified by their OIDs (e.g. OID #1 and OID #2 in field 630).

In an alternative embodiment, the information about the related items within the data set can be stored as an XML document with a document type descriptor (DTD) rather than in database tables. One example of a DTD for such an XML document could be:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<!DOCTYPE affinityChart [
```

```
<!ELEMENT SIMILAR (ITEM,(LIST)*)>
<!ELEMENT LIST (ITEM)+>
<!ATTLIST LIST arm (upper|lower) "upper">
<!ELEMENT ITEM (SEARCH,NAVIGATION)>
<!ATTLIST ITEM type (principal|related) "related">
<!ELEMENT SEARCH (#PCDATA)>
<!ELEMENT NAVIGATION (#PCDATA)>
]>
```

Referring now to Fig. 7, there is shown a visualization of a data set of interrelated items, and further showing the most closely related items to a principal item 711, according to an embodiment of the present invention. In particular, Fig. 7 shows an affinity chart 709 including an affinity curve 710 having a principal item 711 and a plurality of related items 712. Each related item 712 can includes a navigational link 713 and a search link 714. Each navigational link 713 can permit a user to generate a new visualization of a data set of interrelated items related to its associated related item 712. Each search link 714 can permit a user to produce a page that has search results or additional information about the selected item. Adjacent to the principal item 711 and at one end of a selected string of related items 712 is a first sequence element 715, and adjacent to a last item of the selected string of related items 712 is a second sequence element 716, each of which provide a link to a supplementary affinity chart for either more or less remotely relevant strings of related items 712. In Fig. 7, principal item 711 is shown as "union carbide productions". This item represents a musical group. Related musical groups having varying levels of affinity to "union carbide productions" are shown as well. For example, while the related item 712 representing the musical entity "Sarah McLachlan" has an affinity to principal item 711, other related items (such as related item 720 representing the musical entity "the Cure" and related item 725 representing "R.E.M.") have a stronger affinity, as indicated by their closer proximity to principal item 711 on the curve. Similarly, still other related items may

have a weaker affinity to principal item 711, as indicated by their further proximity from principal item 711 on the curve.

Referring now to Fig. 8, there is shown a diagram of other related items to the principal item in Fig. 7, according to the present invention. In particular, Fig. 8 shows an affinity chart 819 representing those items that would have been reached as a result of selecting first sequence element 715 in Fig. 7. In particular, Fig. 8 shows an affinity chart 819 including an affinity curve 820 including a principal item 711 and a plurality of related items 822. Each related item 822 includes a navigational link 823 and a search link 824. Adjacent to the principal item 711 and at one end of a selected string of related items 822 is a first sequence element 825, which provides a link to the affinity chart of more strongly related items, and adjacent to a last item of the selected string of related items 822 is a second sequence element 826 which provides a link to a supplementary affinity chart for more remotely relevant strings of related items 822. By selecting the first sequence element 826, the user can navigate to reach the affinity curve 710 to view a string of relevant items as shown in Fig. 7.

The methods and apparatuses of the present invention provide a visualization technique that allow a set of related data items to be represented with respect to their relationship to a principal data item. The set of related data items may be in any of several forms and the display of the related items eliminates duplication and information overload for the user. The related items may be displayed in many ways, including along curved segments or in list form, so as to allow convenient visualization of the data and the data's relationship to the principal item.

While the invention has been described in detail, including references to specific embodiments, it will be apparent to one skilled in the art that changes and modifications can be made to the invention without departing from the spirit and scope thereof. For example, while a particular embodiment related to musical artists has been disclosed, the invention can be applied equally as well for other data types. Thus, it is intended that the present invention cover the modifications and variations of this invention provided they come within the scope of the appended claims and their equivalents.

*What is claimed is:*

1. A method of information structuring in a data set containing a plurality of interrelated objects, comprising:

ranking related objects based upon relationship strength;

clustering related objects; and

computing the number of affinity charts per object.

2. A method, as in claim 1, wherein said ranking step further comprises:

for each related object to a selected object, calculating an affinity value between each of said related objects and said selected object based upon one or more criteria; and

ordering each of said related objects in said data set according to said affinity value between said related object and said selected object.

3. A method, as in claim 2, wherein said one or more criteria includes an objective measurement.

4. A method, as in claim 2, wherein said one or more criteria includes a subjective measurement.

5. A method of generating a graphical layout, comprising:

selecting a principal node for said graphical layout;

generating at least one affinity chart in connection with said principal node; and

sequentially establishing related items along the at least one affinity chart by rank.

6. A method, as in claim 5, wherein said at least one affinity chart further comprises an affinity curve.

7. A method, as in claim 5, wherein said at least one affinity chart further

comprises a list of related items.

8. The method according to claim 5 further comprising the step of positioning the selected principal node at a prominent location in said graphical layout.

9. The method according to claim 5 wherein for each related item in an affinity chart the size of the item is computed.

10. The method according to claim 6 wherein gradients are used to suggest item affinity level.

11. The method according to claim 10, wherein said gradient further comprises a color gradient.

12. The method according to claim 10, wherein said gradient further comprises a size gradient.

13. The method according to claim 6 wherein provision is made to allow a sufficient vertical and horizontal displacement interval to prevent overlap of related items.

14. A method of information linking, comprising:
hyperlinking related graphs of visualizable large interrelated data sets; and
navigating between hyperlinked related graphs.

15. A method for providing visualization of items from data sets, the method comprising:
determining, for a plurality of items from said data set, a set of properties, said set of properties including a relationship to each other of the subsets of items in the data set, and a value applied to the relationships between the items;

applying local rankings of the relationships between terms, by ranking items i that relate to each item j, and ranking all items k to which item j relates, thereby ranking the affinity of each item j to item sets i and k;

generating a visualization by presenting results separately for each item in a data set and adjusting the presentation to avoid information overlap and overload; and

providing separate presentation for each item of the data set by generating an affinity chart for each item j in the data set, thereby displaying items closely related to selected item j, with item j placed prominently in the affinity chart, and placing items which are more strongly related to j closer to j.

16. The method of claim 15, further comprising the step of expressing closeness along shaped segments, emanating from j's position.

17. A method as in claim 16, wherein said shaped segments further comprise curved segments.

18. The method of claim 15, further comprising:

employing continuous curves including spiral segments, in order to connect items relating to a item (j) at different intensity levels;

adjusting the visualization to avoid information overlap and overload, the items related to item (j) grouped by strength of affinity;

providing an affinity chart, and spacing each related item individually with each item placed in a non-overlapping position;

presenting items with large numbers of related items with multiple affinity charts, and in the case of multiple affinity charts, providing a first affinity chart to visually represent a set of most strongly related items and providing next or subsequent related affinity charts to visually represent less strongly related items;

using curves to represent a relationship of items related to a particular item positioned at a starting point for the curve, with distance along the curve representing a strength of an affinity to the item at the starting point of the curve; and

selectively employing color and shading gradations and curve thickness gradations are to emphasize the curve's role in conveying affinity strength, while placing items so they do not overlap or crowd each other.

19. A method for providing visualization of data sets containing a large number of items from said data sets, the method comprising:

employing continuous curves including spiral segments, in order to connect items relating to a primary item at different intensity levels;

adjusting the visualization to avoid information overlap and overload, the items related to said primary item grouped by strength of affinity;

providing an affinity chart, and spacing each related item individually with each item placed in a non-overlapping position;

presenting items with large numbers of related items with multiple affinity charts, and in the case of multiple affinity charts, providing a first affinity chart to visually represent a set of most strongly related items and providing next or subsequent related affinity charts to visually represent less strongly related items;

using curves to represent a relationship of items related to a particular item positioned at a starting point for the curve, with distance along the curve representing a strength of an affinity to the item at the starting point of the curve; and

selectively employing color and shading gradations and curve thickness gradations are to emphasize the curve's role in conveying affinity strength, while placing items so they do not overlap or crowd each other.

20. A method for providing visualization of arbitrarily large data sets using low and local computational resources, the method comprising:

determining, for at least a plurality of said data sets, a set of properties, said set of properties including a relationship to each other of the subsets of items in the data set, and a value applied to the relationships between the items;

determining at least one primary item for the visualization;

applying local rankings of the relationships between terms, by ranking a first relational set of items that relate to the primary item, and ranking a second relational set

of items to which the primary item relates, thereby ranking the affinity to each primary item to the first relations set of items and to the second relational set of items;

generating a visualization by presenting results separately for each item in a predetermined data set and adjusting the presentation to avoid information overlap and overload;

providing separate presentation for each item of the data set by generating an affinity chart for each primary item in the data set, thereby displaying items closely related to a selected primary item, with the primary item placed prominently in the affinity chart, and placing items which are more strongly related to the primary item closer to the primary item;

expressing closeness along curves or shaped segments, connected or emanating from the primary item's position;

said expression of closeness including completely or partially straight shaped segments;

employing continuous curves including spiral segments, in order to connect items relating to a primary item at different intensity levels;

adjusting the visualization to avoid information overlap and overload, the items related to the primary item grouped by strength of affinity;

providing an affinity chart, and spacing each related item individually with each item placed in a non-overlapping position;

presenting items with large numbers of related items with multiple affinity charts, and in the case of multiple affinity charts, providing a first affinity chart to visually represent a set of most strongly related items and providing next or subsequent related affinity charts to visually represent less strongly related items;

using curves to represent a relationship of items related to a particular item positioned at a starting point for the curve, with distance along the curve representing a strength of an affinity to the item at the starting point of the curve; and

selectively employing color and shading gradations and curve thickness gradations are to emphasize the curve's role in conveying affinity strength, while placing items so they do not overlap or crowd each other.

21. A method for providing visualization of large interrelated data sets, the method comprising:

determining a relationship strength of related items in a data set;

for each item in the data set, ranking related items based on the relationship strength;

clustering related items based on said ranking;

computing a number of affinity charts per item;

establishing clusters of related items;

repeating the steps of ranking related items based on the relationship strength and computing the affinity charts until a desired information structure is achieved;

positioning a principal node prominently in the affinity chart; and

generating entries in said affinity chart emanating from the principal node for each of said clusters of related items.

22. A method as in claim 21, wherein said step of generating entries in said affinity chart further comprises laying out graphs of one per cluster of related items.

23. A method as in claim 21, wherein said step of generating entries in said affinity chart further comprises populating a list of related items.

24. A chart server for creating a graphical layout of a visualization of an interrelated data set, comprising:

means for selecting a principal node for said graphical layout;

means for generating at least one affinity curve from an affinity chart in connection with said principal node;

means for sequentially establishing related items along the at least one affinity curve by rank;

means for linking related affinity charts of interrelated data sets; and

means for navigating between linked related affinity charts.

25. An affinity server for structuring information for use in visualization of a large

interrelated data set, comprising:

means for ranking related objects based upon relationship strength;

means for clustering related objects; and

means for computing the number of affinity charts per object.

26. A computer readable medium containing computer program instructions for providing visualization of items from data sets, said computer program instructions containing instructions for:

determining, for at least a plurality of said data sets, a set of properties, said set of properties including a relationship to each other of the subsets of items in the data set, and a value applied to the relationships between the items;

applying local rankings of the relationships between terms, by ranking items i that relate to each item j, and ranking all items k to which item j relates, thereby ranking the affinity to each item j to item sets i and k;

generating a visualization by presenting results separately for each item in a predetermined data set and adjusting the presentation to avoid information overlap and overload; and

providing separate presentation for each item of the data set by generating an affinity chart for each item j in the data set, thereby displaying items closely related to selected item j, with item j placed prominently in the affinity chart, and placing items which are more strongly related to j closer to j.

27. A system for providing visualization of items from data sets at a first computer operably coupled to a second computer over a communications network, comprising:

a computerized server associated with said second computer, said computerized server including data set visualization software executable on said computerized server and configured to:

determine, for a plurality of said data sets, a set of properties, said set of properties including a relationship to each other of the subsets of items in the data set, and a value applied to the relationships between the items;

apply local rankings of the relationships between terms, by ranking items i that relate to each item j, and ranking all items k to which item j relates, thereby ranking the affinity of each item j to item sets i and k;

generate a visualization by presenting results separately for each item in a data set and adjusting the presentation to avoid information overlap and overload; and

providing separate presentation for each item of the data set by generating an affinity chart for each item j in the data set, thereby displaying items closely related to selected item j, with item j placed prominently in the affinity chart, and placing items which are more strongly related to j closer to j.

1/8

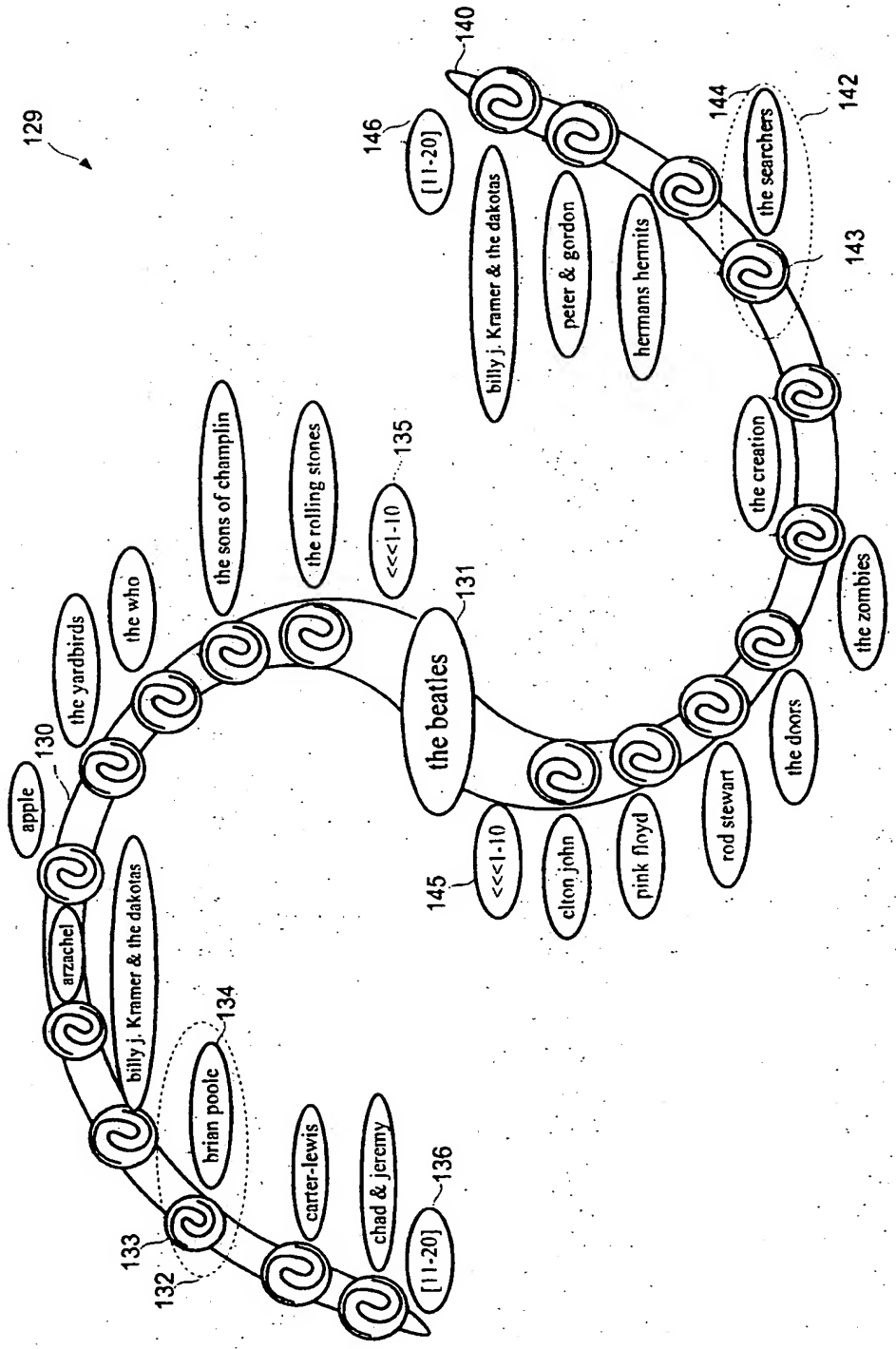
129
140
146
[11-20]
billy j. kramer & the dakotas
peter & gordon
hermans hermits
144
the searchers
142
143
the creation
the zombies
the doors
rod stewart
pink floyd
elton john
<<<1-10
145
131
the beatles
135
<<<1-10
the rolling stones
the sons of champlin
the who
the yardbirds
130
apple
arzachel
billy j. kramer & the dakotas
134
brian poole
carter-lewis
chad & jeremy
136
[11-20]
133
132

Fig. 1

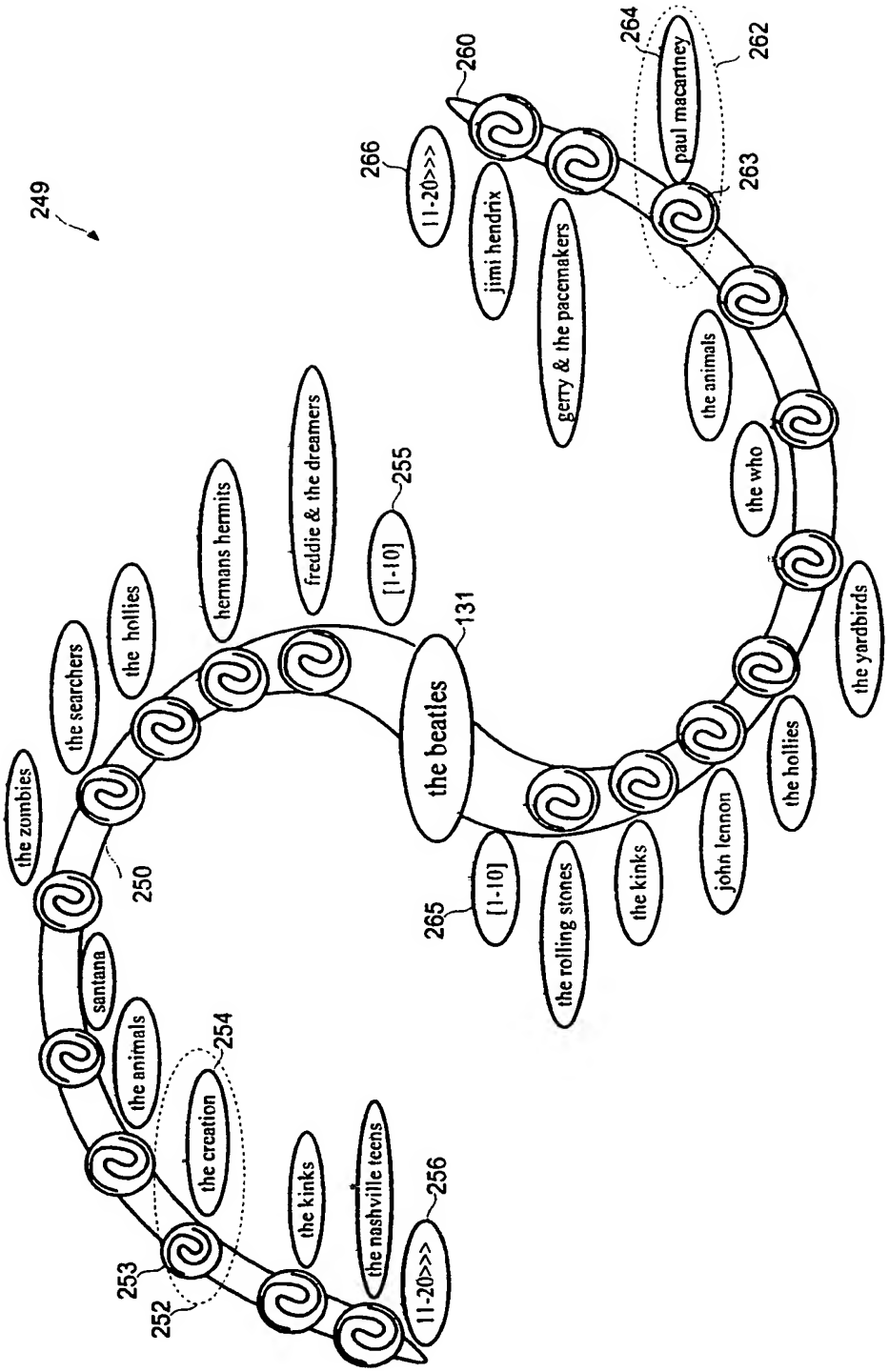
249
250
252
253
254
255
256
260
262
263
264
265
266
the zombies
the searchers
the hollies
hermans hermits
freddie & the dreamers
[1-10]
the beatles
131
[1-10]
the rolling stones
the kinks
john lennon
the hollies
the yardbirds
the who
the animals
gerry & the pacemakers
jimi hendrix
11-20>>>
paul macartney
santana
the animals
the creation
the kinks
the nashville teens
11-20>>>

Fig. 2

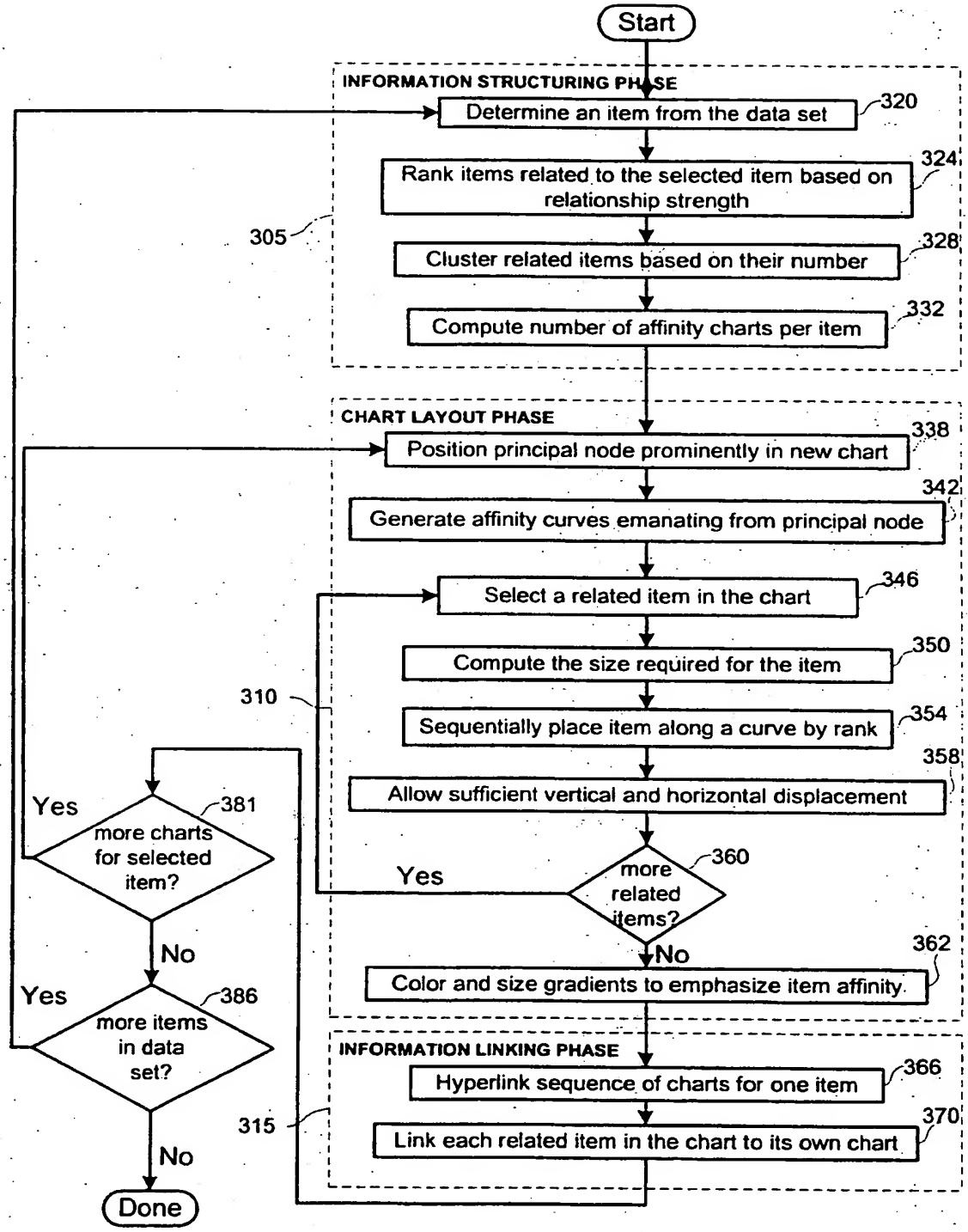
Start
INFORMATION STRUCTURING PHASE
Determine an item from the data set
320
Rank items related to the selected item based on relationship strength
324
305
Cluster related items based on their number
328
Compute number of affinity charts per item
332
CHART LAYOUT PHASE
Position principal node prominently in new chart
338
Generate affinity curves emanating from principal node
342
Select a related item in the chart
346
Compute the size required for the item
350
310
Sequentially place item along a curve by rank
354
Allow sufficient vertical and horizontal displacement
358
more related items?
360
Yes
No
Color and size gradients to emphasize item affinity
362
INFORMATION LINKING PHASE
Hyperlink sequence of charts for one item
366
315
Link each related item in the chart to its own chart
370
more charts for selected item?
381
Yes
No
more items in data set?
386
Yes
No
Done

Fig. 3

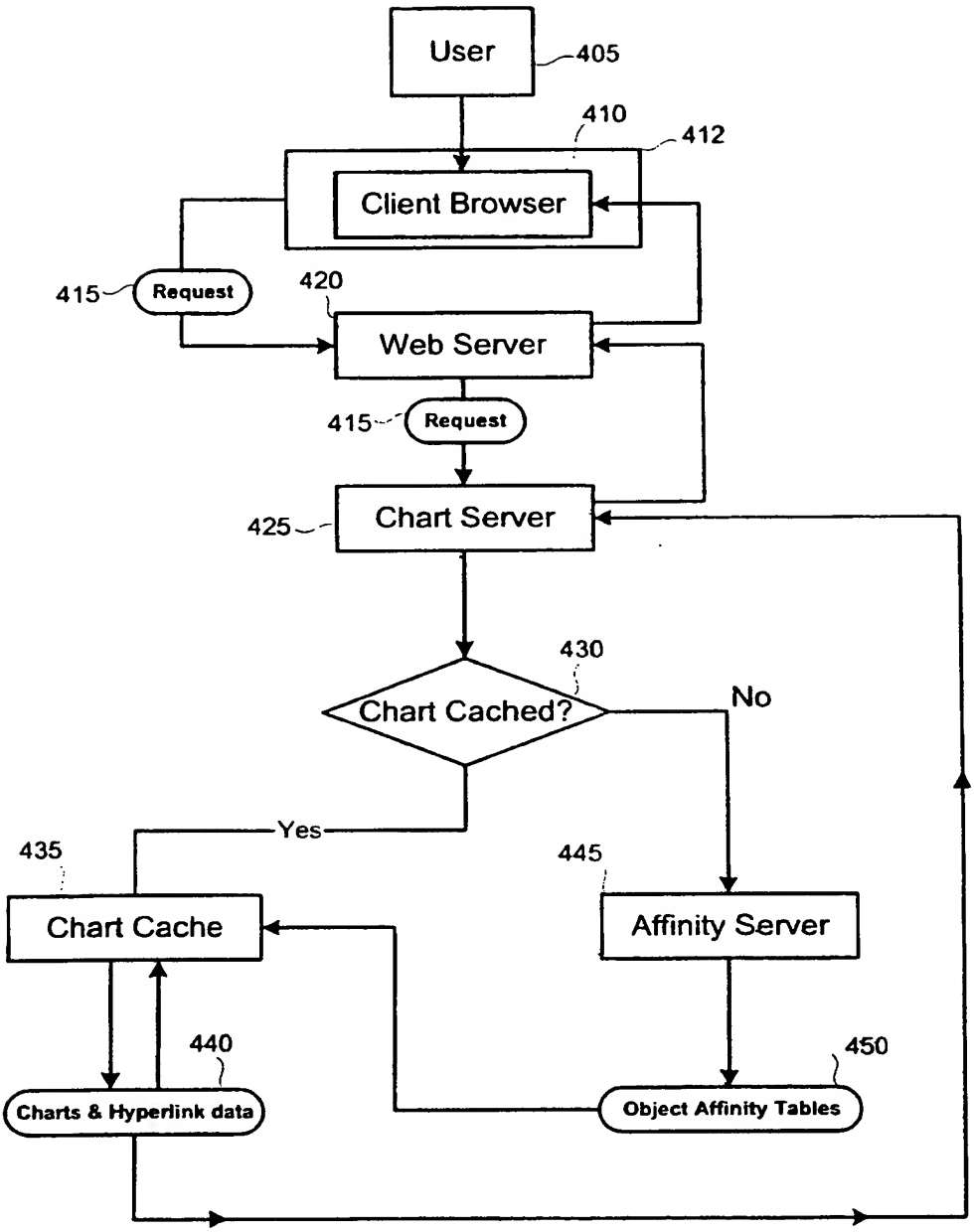
User
405
410
412
Client Browser
415
Request
420
Web Server
415
Request
425
Chart Server
430
Chart Cached?
No
Yes
435
Chart Cache
445
Affinity Server
440
450
Charts & Hyperlink data
Object Affinity Tables

Fig. 4

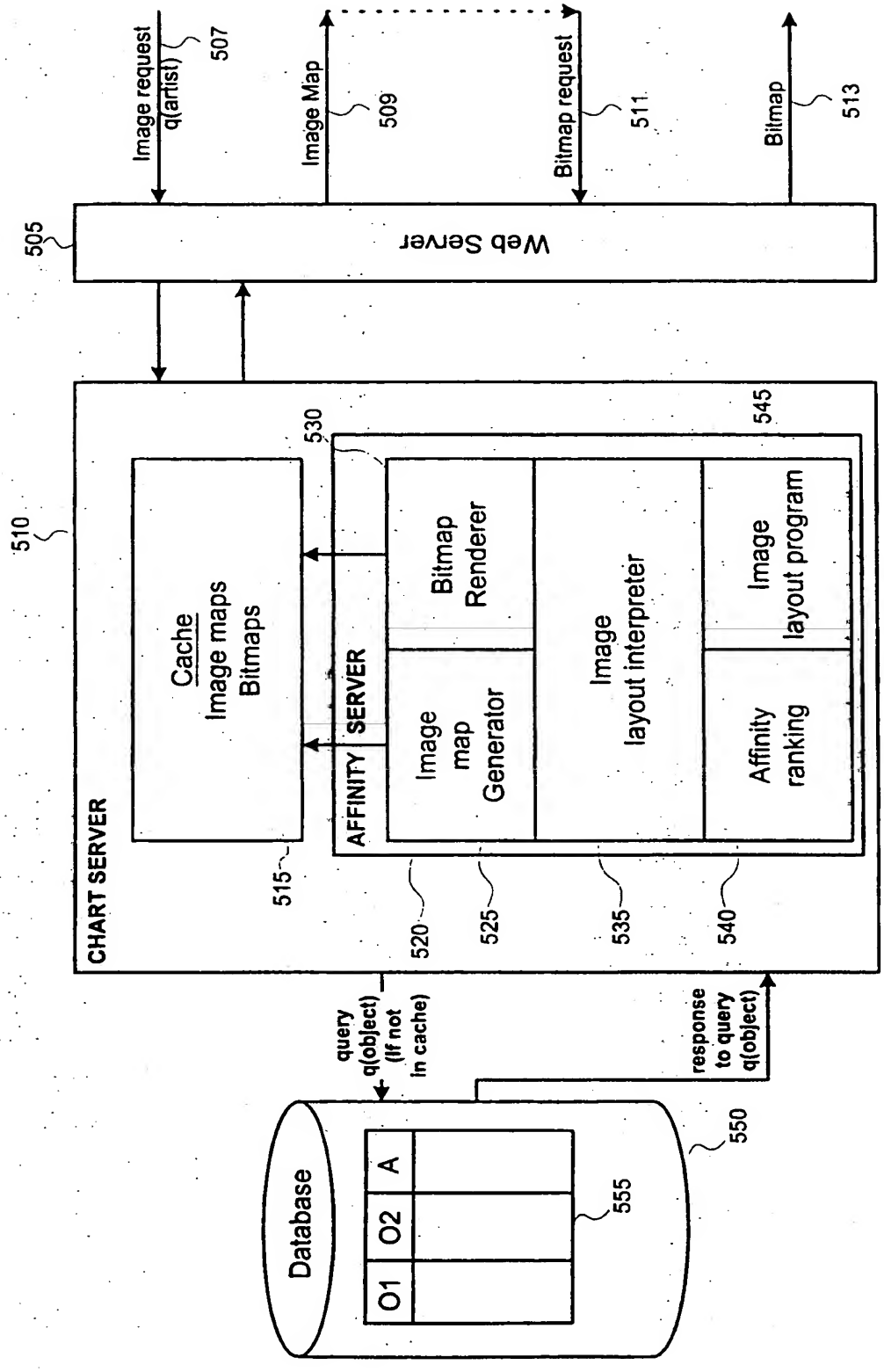
Image request q(artist)
507
Image Map
509
Bitmap request
511
Bitmap
513
Web Server
505
510
CHART SERVER
Cache
Image maps
Bitmaps
515
530
AFFINITY SERVER
Image map Generator
Bitmap Renderer
545
Image layout interpreter
Affinity ranking
Image layout program
520
525
535
540
query q(object) (If not in cache)
response to query q(object)
Database
O1
O2
A
555
550

Fig. 5

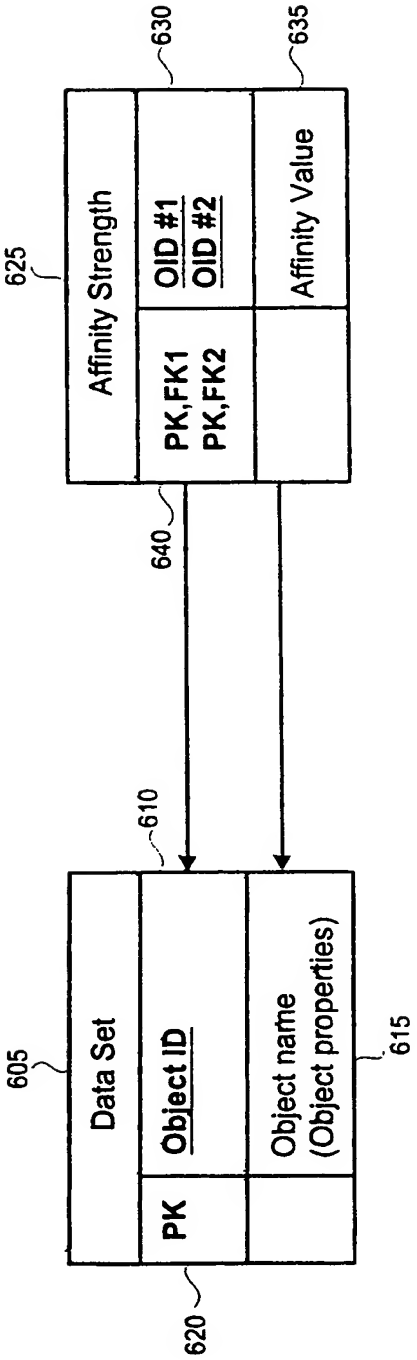
630
635
625
Affinity Strength
OID #1
OID #2
Affinity Value
PK,FK1
PK,FK2
640
610
605
Data Set
Object ID
Object name
(Object properties)
615
PK
620

Fig. 6

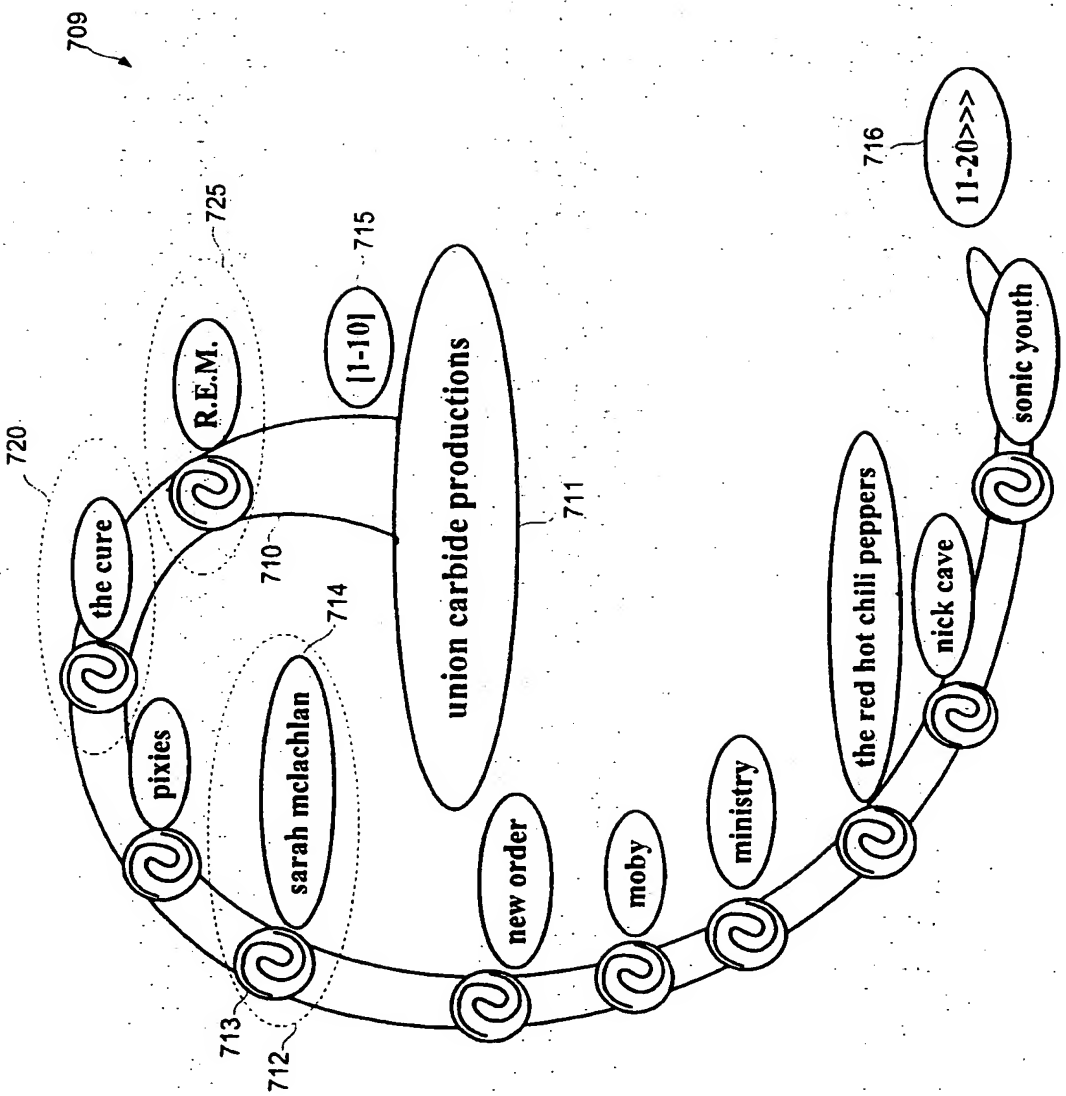
709
716
11-20>>>
725
715
[1-10]
union carbide productions
R.E.M.
sonic youth
720
711
the cure
710
nick cave
714
the red hot chili peppers
sarah mclachlan
pixies
ministry
new order
moby
713
712

Fig. 7

# 8/8

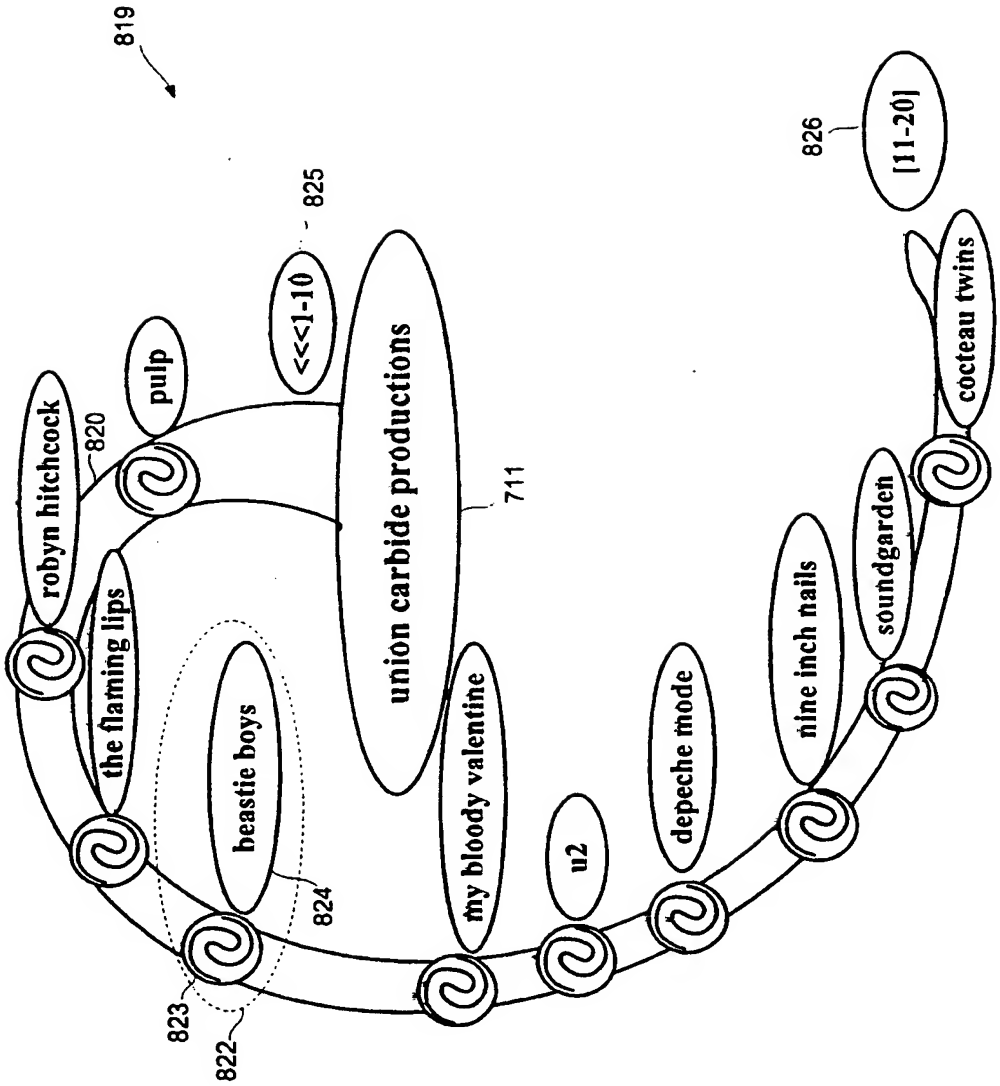
819
826
[11-20]
825
<<<1-10
union carbide productions
711
pulp
820
robyn hitchcock
the flaming lips
beastie boys
824
823
822
my bloody valentine
u2
depeche mode
nine inch nails
soundgarden
cocteau twins

Fig. 8